

Effective Hamiltonians, prethermalization and slow energy absorption in periodically driven many-body systems

Dmitry A. Abanin,¹ Wojciech De Roeck,² and Wen Wei Ho¹

¹*Department of Theoretical Physics, University of Geneva*

²*Instituut voor Theoretische Fysica, KU Leuven, Belgium*

(Dated: October 13, 2015)

We establish some general dynamical properties of lattice many-body systems that are subject to a high-frequency periodic driving. We prove that such systems have a quasi-conserved extensive quantity H_* , which plays the role of an effective static Hamiltonian. The dynamics of the system (e.g., evolution of any local observable) is well-approximated by the evolution with the Hamiltonian H_* up to time τ_* , which is exponentially long in the driving frequency. We further show that the energy absorption rate is exponentially small in the driving frequency. In cases where H_* is ergodic, the driven system prethermalizes to a thermal state described by H_* at intermediate times $t \lesssim \tau_*$, eventually heating up to an infinite-temperature state at times $t \sim \tau_*$. Our results indicate that rapidly driven many-body systems generically exhibit prethermalization and very slow heating. We briefly discuss implications for experiments which realize topological states by periodic driving.

PACS numbers: 73.43.Cd, 05.30.Jp, 37.10.Jk, 71.10.Fd

Introduction. Advances in laser cooling techniques have resulted in experimental realization of well-isolated, highly tunable quantum many-body systems of cold atoms [1]. A rich experimental toolbox of quantum optics available in these systems, combined with the systems' slow intrinsic time scales, allow for a preparation of non-equilibrium many-body states, as well as precise characterization of their quantum evolution. This has made the study of different dynamical regimes in many-body systems one of the forefront directions in modern condensed matter physics (for a review, see [2]).

Conventional wisdom suggests that in a majority of many-body systems, the Hamiltonian evolution starting from a non-equilibrium state should lead to thermalization at sufficiently long times: that is, physical observables reach thermal values, given by the micro-canonical ensemble. Thermalization in such *ergodic* systems is understood in terms of the properties of individual eigenstates, which exhibit thermal observables [3–5] (the eigenstate thermalization hypothesis (ETH)). However, while ETH implies eventual thermalization, it does not make predictions regarding the intermediate-time dynamics of the system. Therefore, much work has been dedicated to studying how thermal equilibrium emerges in different many-body systems.

In particular, there is a class of systems, which exhibit a phenomenon of *prethermalization* [6–9]. Such systems have a set of approximate conservation laws, in addition to energy; therefore, at intermediate time scales they equilibrate to a state given by the generalized Gibbs ensemble (GGE), which is restricted by those conservation laws. Full thermal equilibrium is reached at much longer time scales, set by the relaxation times of the approximate integrals of motion. Prethermalization has been experimentally observed in a nearly integrable one-dimensional Bose gas [10].

In this paper, we establish some general properties of dynamics of periodically driven many-body systems (Floquet systems). Periodic driving is naturally realized in cold atomic systems by applying electromagnetic fields. Periodic driving in quantum systems has recently attracted much theoretical attention, because, amongst many applications, it provides a tool for inducing effective magnetic fields, and for modifying topological properties of Bloch bands [12–18]. For example, one can induce topologically non-trivial Bloch bands (Floquet topological insulators) in non-interacting systems. However, since periodic driving breaks energy conservation, driven ergodic (many-body) systems are expected to heat up, eventually evolving into a featureless, infinite-temperature state [19–22]. Thus, many-body effects are expected to generally make the Floquet topological insulators unstable. Below, we derive general bounds for energy absorption rates in periodically driven many-body systems, which can be applied for instance to understand the lifetimes of Floquet topological insulators.

As the main result of the paper, we show that rapidly driven many-body systems with local interactions generally have a local, quasi-conserved extensive quantity, H_* , which plays the role of an effective Hamiltonian. At times $t \lesssim \tau_*$, the time evolution of any local observable is well-approximated by the Hamiltonian evolution with the time-independent Hamiltonian H_* . Thus, assuming that the Hamiltonian H_* is ergodic, the system exhibits prethermalization to a thermal state described by the Hamiltonian H_* , with an effective temperature set by the initial “energy” $\langle \psi_0 | H_* | \psi_0 \rangle$. The quasi-conservation of H_* is destroyed at timescale $t \sim \tau_*$, when energy absorption occurs and an infinite-temperature state is formed. We show that the heating timescale τ_* is exponentially large in the driving frequency ω :

$$\tau_* \propto e^{C \frac{\omega}{\hbar}}, \quad (1)$$

where C is a numerical constant of order one, and h has the meaning of a maximum energy per particle or spin, and is precisely defined below. Thus, rapidly driven many-body systems generically have a very long prethermalization regime, and absorb energy exponentially slowly. We emphasize that these results are non-perturbative; they generalize and complement our previous work [26], where bounds on linear-response heating rates were proven.

Setup and methods. We consider a quantum system subject to a drive with a period $T = 2\pi/\omega$, described by a time-dependent Hamiltonian:

$$H(t) = H + V(t), \quad V(t+T) = V(t), \quad (2)$$

where H is time-independent, and, without loss of generality, the time-average of the driving term is chosen to be zero, $\int_0^T V(t) dt = 0$. The dynamics of the system is described by the unitary evolution operator $U(t)$, which obeys an equation:

$$i\partial_t U(t) = H(t)U(t), \quad U(0) = I, \quad (3)$$

where I is the identity operator.

Floquet theory (for a review, see [27]) predicts that the solution of Eq.(3) can be written in the following form:

$$U(t) = P(t)e^{-iH_F t}, \quad (4)$$

where $P(t+T) = P(t)$ is a time-periodic unitary such that $P(0) = I$, and H_F is a time-independent Floquet Hamiltonian. In particular, evolution operator over one period is given by:

$$U(T) = \mathcal{T} \exp \left(-i \int_0^T H(t) dt \right) = e^{-iH_F T}. \quad (5)$$

Thus, the evolution of the system at stroboscopic times $t_n = nT$, $n \in \mathbb{Z}$ is governed by the time-independent Hamiltonian H_F . Note that the choice of H_F is not unique: given a particular H_F and projectors $P_i = |i\rangle\langle i|$ onto its eigenstates with energy E_i , the Hamiltonian $H'_F = H_F + \sum_i 2\pi m_i P_i$ is also a valid Floquet Hamiltonian for any $m_i \in \mathbb{Z}$.

Typically, there is no closed form solution of Eq.(4), and one relies on iterative schemes such as the Magnus expansion to obtain H_F for high-frequency drives (for a recent review, see [27, 28]). In this approach, H_F is expanded in terms of powers of T (equivalently, of inverse frequency $1/\omega$), $H_F = \sum_n H_F^{(n)}$, where $H_F^{(n)} = O(T^n)$. The formal solution of Eqs.(4,3) then gives $H_F^{(n)}$ expressed in terms of nested commutators of $H(t)$ at different times. However, the Magnus expansion is only known to converge for bounded Hamiltonians, such that $\|H(t)\|T \leq r_c$, with $r_c \sim 1$, $\forall t$ [28]. Since many-body systems have extensive energies and do not satisfy this

condition, the Magnus expansion is expected not to converge in this case. Indeed, the existence of a quasi-local Floquet Hamiltonian H_F would imply that the system does not heat up to an infinite-temperature state at long times, contrary to the general arguments based on the ETH [21].

Therefore, we propose an alternative approach. The central idea is as follows: we repeatedly transform the Hamiltonian, systematically removing time-dependent terms at increasing order in T . Truncating the procedure at some optimal order n_* (defined below), we obtain a quasi-conserved time-independent Hamiltonian operator H_* . We start by the transformation of the system's wave function $|\psi(t)\rangle$ by a time-periodic unitary $Q(t+T) = Q(t)$, with $Q(0) = I$:

$$|\varphi(t)\rangle = Q(t)|\psi(t)\rangle. \quad (6)$$

Importantly, the wave function $|\varphi(t)\rangle$ coincides with the original wave function $|\psi(t)\rangle$ at stroboscopic times t_n . Its evolution is described by the Schroedinger equation

$$i\partial_t |\varphi(t)\rangle = H'(t)|\varphi(t)\rangle, \quad (7)$$

with a modified Hamiltonian:

$$H'(t) = Q^\dagger H(t)Q - iQ^\dagger \partial_t Q. \quad (8)$$

Thus, the transformation $Q(t)$ defines a new periodic Hamiltonian $H'(t)$, which gives the same stroboscopic evolution as the original Hamiltonian $H(t)$.

For our purposes, it is convenient to write the operator Q as an exponential of a periodic operator $\Omega(t+T) = \Omega(t)$, which is anti-Hermitian, $\Omega^\dagger = -\Omega$, and to represent Ω as an n_{\max} -degree polynomial in the driving period T :

$$Q(t) = e^\Omega, \quad \Omega = \sum_{p=1}^{n_{\max}} \Omega_p, \quad \Omega_p = O(T^p). \quad (9)$$

Using Dunhamel's formula, for $Q = e^\Omega$, Eq.(8) can be rewritten as follows:

$$H'(t) = e^{-\text{ad}_\Omega} (H + V(t)) - i \frac{1 - e^{-\text{ad}_\Omega}}{\text{ad}_\Omega} \partial_t \Omega, \quad (10)$$

where ad_Ω is the adjoint operator, i.e. $\text{ad}_\Omega A = [\Omega, A]$ for any operator A .

We will choose Q such that it minimizes the norm of the time-dependent part of $H'(t)$. For a given n_{\max} , this means that the operators Ω_p are chosen in such a way that the time-dependent part of $H'(t)$, given by $V'(t) = H'(t) - (1/T) \int_0^T H'(t) dt$, has minimal norm. The order of the polynomial n_{\max} should be treated as a parameter to be optimized; below we will see that for a given T and $H(t)$, there exists an optimal $n_{\max} = n_*$, for which the driving term becomes minimal. For a many-body system with local interactions, we find that the optimal $n_* \propto \omega$, and for this n_* operator Ω can be chosen in

such a way that driving term is exponentially reduced, by a factor of $e^{-C\frac{T}{h}}$. The time-independent part H_* of the corresponding Hamiltonian $H'(t)$ then represents a quasi-conserved energy.

We focus on the case of a lattice system with local interactions and locally bounded Hilbert space. The original Hamiltonian H , as well as the driving operator $V(t)$, can be written as a sum of local terms:

$$H = \sum_i H_i, \quad V(t) = \sum_i V_i(t), \quad (11)$$

where i runs over all lattice sites, $i = 1, \dots, N$. We restrict here to one-dimensional systems, but this is not crucial to the method, see also [29]. Here the Hilbert space of possible states of site i is finite-dimensional, as is the case for fermions, spins, as well as hard-core bosons. The locality of the interactions means that each term H_i, V_i acts non-trivially on at most R adjacent sites $i, i+1, \dots, i+R-1$. (e.g., for the nearest-neighbor Heisenberg model, $R = 2$); we refer to R as the range of the operator. Each term H_i, V_i is bounded by a constant interaction strength h :

$$\|H_i\| \leq h, \quad \|V_i(t)\| \leq h. \quad (12)$$

Simple example: single rotating frame transformation. To get some intuition regarding the use of our approach, it is instructive to first consider a familiar example of a transformation Q for $n_{\max} = 1$, i.e., a single rotating frame transformation. For $n_{\max} = 1$, $\Omega = \Omega_1 = O(T)$, and Ω_1 is chosen such that the driving term of order T^0 is eliminated in Eq.(10). Since the zeroth-order contribution in (10) is given by $H'_0 = H + V(t) - i\partial_t\Omega_1$, we obtain Ω_1 :

$$\Omega_1(t) = -i \int_0^t V(t') dt'. \quad (13)$$

Next, we find H' by expanding (10) in powers of $T\text{ad}_{\Omega_1}$:

$$H'(t) = H + \sum_{k=1}^{\infty} H^{(k)}(t), \quad (14)$$

where $H^{(k)}(t)$ is the term of order T^k :

$$H^{(k)}(t) = \frac{(-\text{ad}_{\Omega_1})^k}{k!} H + \frac{k(-\text{ad}_{\Omega_1})^k}{(k+1)!} V(t). \quad (15)$$

To the first order in T , the Hamiltonian H' is given by:

$$H'(t) \approx H + \bar{H}^{(1)} + V^{(1)}(t) + O(T^2),$$

where $\bar{H}^{(k)} = \frac{1}{T} \int_0^T H^{(k)}(t) dt$ is the time-independent part of $H^{(k)}(t)$, and $V^{(k)}(t) = H^{(k)}(t) - \bar{H}^{(k)}$ is the new driving term (with zero time-average). A straightforward calculation shows that $\bar{H}^{(1)} = \frac{T}{2} \int_0^T dt_1 \int_0^{t_1} dt_2 [H(t_1), H(t_2)]$ coincides with the second order of the Magnus expansion. The driving term has been reduced by a factor of T compared to the original driving $V(t)$.

An important feature arising in a many-body system is that the full Hamiltonian H' in Eq.(14) contains non-local terms, because $H^{(k)}(t)$ involves k nested commutators of Ω_1 and $H, V(t)$. However, the norm of the non-local terms is decreasing exponentially with k for sufficiently rapid driving, thus H' is quasi-local. Each term $H^{(k)}(t)$ is extensive and can be written as $H^{(k)}(t) = \sum_i H_i^{(k)}(t)$.

In order to estimate the maximum norm of $H_i^{(k)}(t)$, which we denote by $\|H^{(k)}(t)\|_l \equiv \sup_i \|H_i^{(k)}(t)\|$, we first note the following fact: for any two extensive operators $A = \sum_i A_i, B = \sum_i B_i$ of range R_A, R_B , respectively, such that $\|A\|_l \leq a, \|B\|_l \leq b, C = \text{ad}_A B$ has a range of at most $R_C = R_A + R_B - 1$, and $C = \sum_i C_i$, with norm

$$\|C\|_l \leq 2(R_A + R_B - 1)ab. \quad (16)$$

This is because each operator A_i can commute non-trivially with at most $R_A + R_B - 1$ operators B_j . Repeatedly applying this estimate to the operators $\text{ad}_{\Omega_1}^k H, \text{ad}_{\Omega_1}^k V(t)$ that enter Eq.(15), and using the facts that Ω_1 has range R , and $\|\Omega_1\| \leq hT$ (which follows from Eq.(13)), we obtain:

$$\|H^{(k)}(t)\|_l \leq 2h(2hRT)^k, \quad (17)$$

and the range of $H^{(k)}(t)$ equals $k(R-1) + R$. Thus, even though the Hamiltonian H' is nonlocal, the norm of the long-range terms decays quickly, such that an appropriately defined (see [29]) total (local) norm of $\sum_k H^{(k)}(t)$ is $O(T)$. Therefore, the transformation Ω_1 reduces the amplitude of the time-dependent term by a factor of order T , while at the same time making the Hamiltonian quasi-local, and renormalizing its time-independent part.

General case. Next, we proceed to the general case of $n_{\max} > 1$. $\Omega = \sum_{p=1}^{n_{\max}} \Omega_p$, where $\Omega_p = O(T^p)$ is chosen such that the time-dependent terms in the Hamiltonian H' become of the order $T^{n_{\max}}$. The condition that there are no time-dependent terms of the order $T^q, 1 \leq q \leq n_{\max} - 1$ in H' leads to a set of recursive relations for $\Omega_q(t)$. To derive these relations, we first note that the term of the order q in H' has the following form:

$$H^{(q)}(t) = G^{(q)}(t) - i\partial_t \Omega_{q+1}, \quad (18)$$

where $G^{(q)}(t)$ is expressed in terms of $\Omega_1, \dots, \Omega_q$:

$$G^{(q)}(t) = \sum_{k=1}^q \frac{(-1)^k}{k!} \sum_{\substack{1 \leq i_1, \dots, i_k \leq q \\ i_1 + \dots + i_k = q}} \text{ad}_{\Omega_{i_1}} \dots \text{ad}_{\Omega_{i_k}} H(t) + i \sum_{m=1}^q \sum_{k=1}^{q+1-m} \frac{(-1)^{k+1}}{(k+1)!} \sum_{\substack{1 \leq i_1, \dots, i_k \leq q+1-m \\ i_1 + \dots + i_k = q+1-m}} \text{ad}_{\Omega_{i_1}} \dots \text{ad}_{\Omega_{i_k}} \partial_t \Omega_m. \quad (19)$$

It is convenient to separate $G^{(q)}(t)$ into a time-independent part, $\bar{H}^{(q)}$, and a time-dependent part $V^{(q)}$ with zero average over one period:

$$\bar{H}^{(q)} = \frac{1}{T} \int_0^T G^{(q)}(t) dt, \quad V^{(q)}(t) = G^{(q)}(t) - \bar{H}^{(q)}. \quad (20)$$

Then, for $q \leq n_{\max} - 1$, we eliminate the time-dependent term of the order T^q in H' (see Eq.(18)) by choosing Ω_{q+1} as follows:

$$\Omega_{q+1}(t) = -i \int_0^t V^{(q)}(t') dt'. \quad (21)$$

For $q = 0$, $\Omega_1(t)$ is given by Eq.(13).

Relations (19,20,21) define the transformation Ω which makes the time-dependent terms in the Hamiltonian H' of the order $T^{n_{\max}}$:

$$H' = H + \sum_{q=1}^{n_{\max}-1} \bar{H}^{(q)} + \delta H'(t), \quad \delta H'(t) = O(T^{n_{\max}}). \quad (22)$$

Similar to the simple example of $n_{\max} = 1$ considered above, the time-dependent term, $\delta H'(t)$ becomes nonlocal. As a next step, we should find an optimal $n_{\max} = n_*$ for which $\delta H'(t)$ is minimized. To that end, we first estimate how $\|\Omega_q\|_l$ depends on q . We prove that

$$\|G^{(q)}\|_l \leq C(R)^q q! h(hT)^q, \quad (23)$$

where $C(R) = CR$ with C a combinatorial constant of order 1. The other operators have then derived bounds since $\|\bar{H}^{(q)}\|_l \leq \|G^{(q)}\|_l$, $\|\bar{H}^{(q)}\|_l \leq 2\|G^{(q)}\|_l$. For Ω_q , we have

$$\|\Omega_{q+1}\|_l \leq 2C(R)^q q! (hT)^{q+1}. \quad (24)$$

The $q!$ factor in the above bounds arises because of the many-body nature of the system: $G^{(q)}$ involves q nested commutators of $H(t)$, $V(t)$. We refer to the supplementary material for a proof of these bounds and to [29] for generalizations. Eq.(23) shows that there are two competing effects which control the behaviour of $\|\Omega_q\|_l$: suppression of $\|\Omega_q\|_l$ by a factor of T^q , and its growth due to $q!$. Eventually, the factorial dominates and therefore for $q > \frac{1}{C(R)hT}$ $G^{(q)}$ stops decreasing with q . To minimize the norm of the remaining time-dependent terms (see r.h.s. of Eq.(24)), we choose

$$n_* \approx \frac{1}{C(R)hT}, \quad (25)$$

which allows us to reduce the time-dependent term $\delta H'(t)$ by an exponentially large factor of $e^{n_*} \approx e^{-1/(C(R)hT)}$. This implies that the energy absorption rate is exponentially small.

Next, we spell out the consequences for the evolution of a local observable O , with $\|O\| = 1$. Let $U_{s,t}$ be the evolution generated by H' , which we split as $H'(t) = H_* + V^{(n_*)}(t)$. For simplicity we restrict the discussion to stroboscopic times [30] where $O(t)$ is given by $\mathcal{U}_{0,t}O \equiv U_{0,t}OU_{0,t}^*$ and we compare this with the effective evolution $\mathcal{U}_{*,t}O \equiv e^{itH_*}Oe^{-itH_*}$. We use the Duhamel formula

$$\mathcal{U}_{0,t}O = \mathcal{U}_{*,t}O + i \int_0^t ds \mathcal{U}_{t,s}[V^{(n_*)}(s), \mathcal{U}_{*,s}O]$$

and then unitarity and the Lieb-Robinson bound to control (see [29] for details) the norm of the second term by

$$C(Rn_*\|V^{(n_*)}\|_l) \int_0^t ds (v_{LR}s + Rn_*) \quad (26)$$

with v_{LR} a Lieb-Robinson velocity of H_* that is of order Rh , and Rn_* the range of $V^{(n_*)}$. Therefore, this norm remains small for exponentially long times in $1/T$, as well. To be specific, we see that it remains smaller than one up to times of order $\exp(\frac{1}{2C(R)hT})$, where $1/2$ comes from the fact that (26) grows as t^2 .

Discussion. In this paper, we considered many-body systems subject to a high-frequency periodic driving. We have shown that there is a broad time window, $t \lesssim \tau_*$, in which stroboscopic dynamics of such systems is controlled by an effective time-independent Hamiltonian H_* . We have used a series of “gauge”, time-periodic unitary transformations to effectively reduce the strength of the driving term and to establish the existence of H_* . The advantage of our approach compared to the standard Magnus expansion [27, 28] is that it allows us to control the magnitude of the driving terms after the transformations.

We note that recently Canovi et al. [32] and Bukov et al. [33] discussed prethermalization in *weakly interacting* driven systems. Our results complement these works: we have shown that (rapidly) driven interacting systems *generically* exhibit a broad prethermalization regime, which can be observed in a quench experiment as follows. Let us initially prepare the system in some non-equilibrium state $|\psi\rangle$, and subject it to a rapid periodic

drive. At times $t \lesssim \tau_*$ the system will reach a steady state, in which physical observables have thermal values, $\langle \psi(t) | O | \psi(t) \rangle = \text{Tr}(O\rho)$, where the density matrix $\rho \propto e^{-H_*/T_{\text{eff}}}$, with T_{eff} being the effective temperature set by the energy density of the initial state. Thus, at times $t \lesssim \tau_*$ the system appears as if it is not heating up. The system will absorb energy and relax to a featureless, infinite-temperature state beyond times $t \sim \tau_*$. We expect this phenomenon to be observable in driven system of cold atoms and spins (assuming relaxation of spins due to phonons is slow).

Finally, we briefly discuss the implications of our results for the current efforts to realize topologically non-trivial strongly correlated states (e.g., fractional Chern insulators) in periodically driven systems. Experimentally, one tries to design a drive for which the ground state of an effective time-independent Hamiltonian (usually calculated within low-order Magnus expansion) is topologically non-trivial. A central challenge is to prepare the system in a ground state of the effective Hamiltonian. Since we have shown that the dynamics of the system is controlled by H_* up to exponentially long times, one can envision that the “Floquet fractional Chern insulators” can be prepared as follows. Let us assume that the system can be initially prepared in a (topologically trivial) ground state of the Hamiltonian H . Then, the driving is switched on adiabatically to the value which corresponds to the desired effective Hamiltonian H_* . However, the switching should also be done fast compared to τ_* to avoid energy absorption. Since H and H_* describe different phases, the system will necessarily go through a quantum critical point (QCP), and excitations will be created via a Kibble-Zurek mechanism. The number of excitations can be minimized by designing a non-linear passage through the QCP [34]. We leave a detailed exploration of these ideas for future work.

Acknowledgements. D.A. acknowledges support by Alfred Sloan Foundation. W.D.R. also thanks the DFG (German Research Fund), the Belgian Interuniversity Attraction Pole (P07/18 Dygest), as well as ANR grant JCJC for financial support.

-Note added. While this manuscript was being finalized, a related result [36] appeared, building on [35] (local driving). [36] proves a similar bound for the absorption rate in driven systems using a different approach (namely, studying evolution over one driving period). We note that, in addition to providing bounds on energy absorption, our approach allows us to describe the long-time dynamics of the system.

- tore, Rev. Mod. Phys. 83, 863 (2011).
- [3] J. M. Deutsch, Phys. Rev. A **43**, 2046 (1991).
 - [4] M. Srednicki, Phys. Rev. E **50**, 888 (1994).
 - [5] M. Rigol, V. Dunjko, and M. Olshanii, Nature **452**, 854 (2008).
 - [6] J. Berges, S. Borsányi, and C. Wetterich, Phys. Rev. Lett. **93**, 142002 (2004).
 - [7] M. Eckstein, M. Kollar, and P. Werner, Phys. Rev. Lett. **103**, 056403 (2009).
 - [8] M. Moeckel and S. Kehrein, New J. of Phys. **12**, 055016 (2010).
 - [9] L. Mathey and A. Polkovnikov, Phys. Rev. A **81**, 033605 (2010).
 - [10] M. Gring *et al.*, Science **337**, 1318 (2012).
 - [11] S. Fishman, D. R. Grempel and R. E. Prange, Chaos, Quantum Recurrences, and Anderson Localization. Phys. Rev. Lett. **49**, 509 (1982).
 - [12] M. Aidelsburger, N. Atala, M. Lohse, J. T. Barreiro, B. Paredes, and I. Bloch, Phys. Rev. Lett. **111**, 185301 (2013).
 - [13] G. Jotzu *et al.*, Nature **515**, 237 (2014).
 - [14] M. Aidelsburger *et al.*, Nat. Phys. **11**, 162 (2014).
 - [15] T. Oka and H. Aoki, Phys. Rev. B **79**, 081406 (2009).
 - [16] T. Kitagawa, E. Berg, M. Rudner, and E. Demler, Phys. Rev. B **82**, 235114 (2010).
 - [17] N. H. Lindner, G. Refael, and V. Galitski, Nat. Phys. **7**, 490 (2011).
 - [18] M. S. Rudner, N. H. Lindner, E. Berg, and M. Levin, Phys. Rev. X **3**, 031005 (2013).
 - [19] L. D’Alessio and M. Rigol, Phys. Rev. X **4**, 041048 (2014).
 - [20] A. Lazarides, A. Das, and R. Moessner, Phys. Rev. E **90**, 012110 (2014).
 - [21] P. Ponte, A. Chandran, Z. Papić, and D. A. Abanin, Annals of Physics **353**, 196 (2015).
 - [22] It should be noted that driven many-body localized systems can exhibit dynamical localization in the energy space, avoiding heating to an infinite-temperature state, see Refs. [21, 23–25].
 - [23] A. Lazarides, A. Das, R. Moessner, Phys. Rev. Lett. **115**, 030402 (2015).
 - [24] P. Ponte, Z. Papić, F. Huveneers, and D. A. Abanin, Phys. Rev. Lett. **114**, 140401 (2015).
 - [25] D. A. Abanin, W. De Roeck, and F. Huveneers, arXiv:1412.4752 (2014).
 - [26] D. A. Abanin, W. De Roeck, and F. Huveneers, arXiv:1507.01474 (2015).
 - [27] M. Bukov, L. D’Alessio and A. Polkovnikov, Advances in Physics, Vol. 64, No. 2, 139-226 (2015).
 - [28] S. Blanes, F. Casas, J. A. Oteo, J. Ros, Physics Reports **470**, 151 (2009).
 - [29] D. A. Abanin, W. De Roeck, F. Huveneers, W. Ho, arXiv:1509.05386 (2015).
 - [30] For general times, $e^{-\Omega(t)}O(t)e^{\Omega(t)} = \mathcal{U}_{0,t}O$, leading to the same estimate but with a periodic error of order $1/T$.
 - [31] E.H. Lieb, D. Robinson, Comm. Math. Phys. **28**: 251-257(1972), B. Nachtergaele, R. Sims, arXiv:1004.2086.
 - [32] E. Canovi, M. Kollar, and M. Eckstein, arXiv:1507.00991 (2015)
 - [33] M. Bukov, S. Gopalakrishnan, M. Knap, and E. Demler, arXiv:1507.01946 (2015).
 - [34] R. Barankov, A. Polkovnikov, Phys. Rev. Lett. **101**, 076801 (2008).
 - [35] T. Kuwahara, T. Mori, K. Saito, arXiv:1508.05797, 2015

-
- [1] I. Bloch, J. Dalibard, and W. Zwerger, Rev. Mod. Phys. **80**, 885 (2008).
 - [2] A. Polkovnikov, K. Sengupta, A. Silva, and M. Vengalator

[36] T. Mori, T. Kuwahara, K. Saito, arXiv:1509.03968 (2015).

SUPPLEMENTARY MATERIAL: EFFECTIVE HAMILTONIANS, PRETHERMALIZATION AND SLOW ENERGY ABSORPTION IN PERIODICALLY DRIVEN MANY-BODY SYSTEMS

Here, we provide a proof of the bounds on the terms of the renormalized Hamiltonian H' .

REDUCTION TO COMBINATORIAL PROBLEM

For simplicity, let us list the important bounds on local norms that we claim:

$$\|G^{(q)}\|_l \leq C(R)^q q! h(hT)^q, \quad (\text{S1})$$

$$\|V^{(q)}\|_l \leq 2C(R)^q q! h(hT)^q. \quad (\text{S2})$$

$$\|\Omega_{q+1}\|_l \leq 2C(R)^q q! (hT)^{q+1}. \quad (\text{S3})$$

where, moreover, (S2) and (S3) follow immediately from (S1) for a given q . The operators $V^{(q)}, G^{(q)}, \bar{H}^{(q)}, \Omega_{q+1}$ have range $R(q+1)$. We write $G^0 \equiv H$, consistent with the fact that the range of the original operator H is R . After taking into account the definition of Ω_{q+1} , the relation (19) between these operators is

$$G^{(q)}(t) = \sum_{k=1}^q \frac{(-1)^k}{k!} \sum_{\substack{1 \leq i_1, \dots, i_k \leq q \\ i_1 + \dots + i_k = q}} \text{ad}_{\Omega_{i_1}} \dots \text{ad}_{\Omega_{i_k}} G^{(0)} + i \sum_{m=1}^q \sum_{k=1}^{q+1-m} \frac{(-1)^{k+1}}{(k+1)!} \sum_{\substack{1 \leq i_1, \dots, i_k \leq q+1-m \\ i_1 + \dots + i_k = q+1-m}} \text{ad}_{\Omega_{i_1}} \dots \text{ad}_{\Omega_{i_k}} V^{(m-1)}. \quad (\text{S4})$$

and the only thing in need of a proof is to check the bound on $\|G^{(q)}\|_l$ while assuming the stated bounds on the other operators for all $q' < q$. Bounding all commutators, we see that the local norm of G^q due to the first and second sum are bounded by, respectively

$$a(q) \sum_{n=2} \frac{4^n C(R)^{1-n}}{(n-1)!} \sum_{\{m_j\}}^{q+1} Z_R(\{m_j\}) \chi(m_1 = 1) \prod_{j=1}^n (m_j - 1)! \quad (\text{S5})$$

$$a(q) \sum_{n=2} \frac{4^n C(R)^{1-n}}{n!} \sum_{\{m_j\}}^{q+1} Z_R(\{m_j\}) \prod_{j=1}^n (m_j - 1)! \quad (\text{S6})$$

where we abbreviated

$$a(q) = h(hT)^q C(R)^q,$$

and the sum $\sum_{\{m_j\}}^{q+1}$ is over all sequences $m_j, j = 1, \dots, n$ of positive integers $m_j \geq 1$ satisfying

$$\sum_j m_j = q + 1$$

and $Z_R(\{m_j\})$ is the number of sequences of (discrete) intervals $I_j \subset \mathbb{Z}$ such that

- $|I_j| = Rm_j$ with $|I|$ the number of elements in I .
- For $j > 1$, $I_j \cap \left(\bigcup_{i=1}^{j-1} I_i\right)$ is nonempty.
- $\min I_1 = 1$.

The intersection condition of course derives from the structure of nested commutators. To conclude the proof of the bounds (S1), we have to show that, for some (q -independent) choice of $C(R)$, the sum of these two sums is bounded by $q!$. It is sufficient to prove a bound on (S6), as (S5) reduces to that case upon increasing $C(R) \rightarrow C_1 C(R)$, with C_1 such that $\frac{C_1^{1-n}}{(n-1)!} \leq 1/n!$. Furthermore, it suffices to consider the case $R = 1$. Indeed, for any given sequence $\{m_j\}$, one checks that

$$Z_R(\{m_j\}) \leq (2R)^{n-1} Z_1(\{m_j\})$$

Hence we are down to showing

Lemma 1. *For some C_0 independent of q ,*

$$\sum_{n=2} \frac{C_0^{-n}}{n!} \sum_{\{m_j\}}^{q+1} Z_1(\{m_j\}) \prod_{j=1}^n (m_j - 1)! \leq q!, \quad (\text{S7})$$

The remainder of this supplementary material is devoted to the proof of this lemma, split over 3 sections.

DEFINITIONS AND BASIC BOUND

\mathcal{L} and \mathcal{F} intervals

For a sequence of intervals I_j as above, we classify each interval I as an \mathcal{L} -interval or an \mathcal{F} -interval: I_j is an \mathcal{F} -interval iff. $I_j \subset \cup_{i=1}^{j-1} I_i$, and I_1 is always an \mathcal{L} interval. We call l and f the number of, respectively, \mathcal{L} and \mathcal{F} intervals. It follows that $n = l + f$.

Every \mathcal{L} -interval has at least one site of overlap with the earlier support $\cup_{i=1}^{j-1} I_i$. For each \mathcal{L} -interval I_j we write its length as

$$L_j + A_j, \quad L_j = |I_j \setminus (\cup_{i=1}^{j-1} I_i)| + 1,$$

i.e. such that $A_j + 1$ is the size of the intersection. Note that obviously $A_1 = 0$. We call the length (size) of each

\mathcal{F} interval F_j . Let us now define

$$L = \sum_j L_j - (l - 1), \quad F = \sum_j F_j, \quad A = \sum_j A_j$$

The number L has an important interpretation. It is the size of total set $\cup_{j=1}^n I_j$. It is convenient to also define $G = A + F$. We have the relation (sum of all lengths)

$$q + 1 = L + (l - 1) + G \quad (\text{S8})$$

The most important application of this classification of intervals will be that the number of ways to attach an \mathcal{F} -interval is maximally L (instead of the sum of lengths of all previous integrals, as a more brutal bound would give). The number of ways to attach a new \mathcal{L} -interval is even smaller because one has to place at at the boundary of the total set present up to then.

Basic representation

Our basic bound for the left hand side of (S7) is then

$$w(q) := \sum_l w(q, l) = \sum_{f, l, L, G} C_0^{-(l+f)} \frac{2^{l+1}}{(l+f)!} \frac{(l+f)!}{l!f!} L^G \sum_{\{L_j\}|L, l} W_{\mathcal{L}} \sum_{\{F_j\}|L, f} L^{-F} W_{\mathcal{F}} \sum_{\{A_j\}|\{L_j\}} L^{-A} W_{\mathcal{A}} \quad (\text{S9})$$

where we introduced the weights

$$W_{\mathcal{L}} = \prod_j (L_j - 1)!$$

$$W_{\mathcal{F}} = L^f \prod_j (F_j - 1)!$$

$$W_{\mathcal{A}} = \prod_j \frac{(L_j + A_j - 1)!}{(L_j - 1)!}$$

and a notation like $\sum_{\{L_j\}|L, l}$ clarifies that the sum is performed with L, l held fixed. The respective sums run over all sequences of lengths L_j, F_j, A_j . Here, and it what follows, we do not indicate explicitly constraints that are obvious from the definitions, e.g. $l < L$ and the sum rule (S8). Let us clarify (S9). The factor $\frac{(l+f)!}{l!f!}$ is the number of ways we can divide n elements into l and f elements, i.e. to choose for which $j = 1, \dots, n$ the \mathcal{L} intervals occur. The number 2^{l+1} comes from choosing for each \mathcal{L} -interval (but not the first), whether it explores the lattice to the left or the right of the preexisting intervals. The factor $(l+f)!$ is the denominator is simply the original factor $n!$. The factor L^f in $W_{\mathcal{F}}$ bounds the number of ways the f intervals \mathcal{F} can be placed in a stretch of length

L . Finally, recall that the local norm refers to splitting $G^{(q)} = \sum_i G_i^{(q)}$ where each G_i acts on the sites $i, i+1, \dots, i+R(q+1)-1$ (since $R(q+1)$ is the range of $G^{(q)}$, i.e.

$$\|G^{(q)}\|_l = \sup_i \|G_i^{(q)}\|.$$

In contrast, we sum the norms of all terms that are 'grown' out of the origin (I_1 starts at 1) and hence can act on any stretch of sites of length $R(q+1)$ containing the origin. By reordering terms, and noting that we are anyhow working with quantities that are suprema over the lattice, we recover the local norm.

The main point is that the number of possibilities for the \mathcal{F} intervals is restricted by L . This provides a competition between G and L since their sum is constraint by (S8).

AUXILIARY LEMMA'S

We state some straightforward bounds on sums of $W_{\mathcal{L}}, W_{\mathcal{A}}, W_{\mathcal{F}}$.

Lemma 2.

$$\sum_{\{F_j\}|L, f} L^{-F} W_{\mathcal{F}} \leq (1 + \mathcal{O}(1/L))^f$$

Proof. We distribute the factor L^{-F} over the blocks

$$\sum_{\{F_j\}|L,f} L^{-F} W_{\mathcal{F}} \leq \prod_j \sum_{F_j} L^{1-F_j}$$

and each product is bounded as

$$\sum_{1 \leq F_j \leq L} L^{1-F_j} = 1 + \mathcal{O}(1/L)$$

this yields the lemma. \square

Lemma 3.

$$\sum_{\{A_j\}|\{L_j\}} L^{-A} W_{\mathcal{A}} \leq C^l \sqrt{L}$$

Proof. To control this we estimate

$$\sum_{0 \leq A_j \leq L-L_j} L^{-A_j} \frac{(L_j + A_j - 1)!}{(L_j - 1)!} \leq \sum_{0 \leq A_j \leq L-L_j} \prod_{n=0}^{A_j} \frac{L_j + n}{L}$$

and then we take the product over j . If $L_j/L \leq a < 1$, then obviously this expression is $1 + C(a)/L$. In case we have no good upper bound on L_j/L , then we have the more brutal bound

$$\sum_{0 \leq A_j \leq L-L_j} \prod_{n=0}^{A_j} \frac{L_j + n}{L} \leq \sum_{0 \leq A_j \leq L-L_j} \left(\frac{L - A_j/3}{L} \right)^{A_j/3}$$

which is bounded by $C\sqrt{L}$. Since $\sum_{j=1}^l L_j = L + l - 1$ and each $L_j \geq 2$, there can be at most one L_j for which, say $L_j > 0.9L$. Thus the brutal bound is invoked at most once and this yields the lemma. \square

Lemma 4.

$$\sum_{\{L_j\}|L,l} W_{\mathcal{L}} \leq C^l \frac{(L-1)!}{L^{l-1}}$$

Proof. By Stirling and $\sum_j (L_j - 1) = L - 1$, we have

$$\frac{W_{\mathcal{L}}}{(L-1)!} \leq \frac{C^l}{\sqrt{L}} \left(\prod_j \left(\frac{L_j - 1}{L} \right)^{L_j - 1} \sqrt{L_j - 1} \right).$$

We note that $L_j \geq 2$. The leading contribution to the product comes from $\{L_j\}$ where all but one of the L_j is at its minimal value 2, and hence the remaining L_j is not smaller than $L - 2l$. All other contributions are down by powers of L and can be summed. \square

CONCLUSION OF PROOF

We put the above lemmata together, getting

$$\frac{w(q, l)}{q!} \leq \sum_{f, L, G} \frac{C^{l+f}}{l! f!} \frac{(L-1)! L^G}{q!} \left(\frac{1}{L} \right)^{l-3/2} \quad (\text{S10})$$

and we have to show that the sum over l of this is bounded by a constant. We will eventually use the bound (S10) only to estimate the $l > 1$ terms, whereas for the $l = 1$ terms we need a separate argument that is a bit less wasteful. Let us abbreviate $k = l - 1$, then, by using the sum rule $L + G = q + 1 - k$ and Stirling, we get

$$\frac{(L-1)! L^G}{q!} \leq C q^{-k} e^{G+k} \left(\frac{q - G - k}{q} \right)^{q-k} e^{\frac{2G}{q-G-k}}$$

Plugging into (S10) and restricting to $l > 1$, hence $k > 0$, we can now bound all sums by a constant.

This just leaves

The case $l = 1$

If $l = 1$, then $A = 0$, hence $F = G = q + 1 - L \geq 1$. We proceed as in the buildup to (S10) but we note that the factor \sqrt{L} from the sum over $\{A_j\}$ can be dropped for $l = 1$. We get then

$$\frac{w(q, 1)}{q!} \sum_{f, L, F} \frac{C^f}{f! (r-1)!} (L-1)! L^F$$

At this point, we recall the relationship between F and f . Looking back at Lemma 2, we see that we can easily extract an extra exponential decay factor, say $Ce^{-(F-2f)}$ forcing F to be close to f . With this improvement, we can sum up the contributions from $l = 1$ as well.